THE UNIVERSITY OF NORTH CAROLINA RECORD

MAY, 1917
NUMBER 147

MEASUREMENT OF ACHIEVEMENT IN THE FUNDAMENTAL ELEMENTARY SCHOOL SUBJECTS

FACULTY COMMITTEE ON EXTENSION

LOUIS R. WILSON M. C. S. NOBLE COLLIER COBB E. C. BRANSON
H. W. CHASE L. A. WILLIAMS E. R. RANKIN A. H. PATTERSON
N. W. WALKER M. H. STACY C. L. RAPER

RALEIGH
EDWARDS & BROUGHTON PRINTING COMPANY
1917

The School of Education of the University of North Carolina

FACULTY

EDWARD KIDDER GRAHAM, A.M., D.C.L., LL.D., *President.*

MARCUS CICERO STEPHENS NOBLE, *Professor of Pedagogy and Dean of the School of Education.*

NATHAN WILSON WALKER, A.B., *Professor of Secondary Education and State Inspector of High Schools.*

HARRY WOODBURN CHASE, Ph.D., *Professor of Psychology.*

LOUIS ROUND WILSON, Ph.D., *Professor of Library Administration.*

LESTER ALONZO WILLIAMS, Pd.D., *Professor of School Administration.*

JOSEPH HENRY JOHNSTON, Ph.D., *Assistant Professor of School Administration.*

COURSES LEADING TO THE DEGREE OF BACHELOR OF ARTS IN EDUCATION

To intending teachers the School of Education offers a course of study extending over four years and leading to the degree of Bachelor of Arts in Education. The professional courses required include Secondary School Organization and Administration, Educational Psychology, North Carolina School Conditions, and Supervision of Teaching. Opportunity for elective work in the History of Education and School Administration is afforded. Systematic observation of teaching in the town school is carried on in several of the courses. Laboratory facilities for work in Educational Psychology are available.

COURSES LEADING TO THE DEGREE OF MASTER OF ARTS

Both in the regular and in the summer session, the School of Education offers work for graduate students who are candidates for the degree Master of Arts. In order to be granted this degree, candidates who major in Education must conform to the usual requirements for admission to the Graduate School, and must complete satisfactorily a total of fifteen hours work of an advanced grade, ten hours of which must be chosen from courses in the Graduate School, and eight hours of which must be in Education. A satisfactory thesis on an approved subject must be presented before the degree will be granted.

For catalogue and further information, address

THOMAS J. WILSON, JR.,
Registrar.

INTRODUCTION

There is manifest in the State an increasing interest in various phases of educational measurement. Judging from the inquiries coming to the School of Education, there seems a place for a statement in simple form of the nature and value of some of the more common standard tests and scales. The bulletin which follows is simply an introduction to the field of work with which it deals, and its main function should be to stimulate to wider reading and further work.

The School of Education will be glad to receive statements of the results attained by any one in giving these tests, and to assist in their interpretation. It is hoped that by the coöperation of school officials State-wide standards may gradually be built up and published for the benefit of those engaged in school work.

H. W. Chase.

Digitized by the Internet Archive
in 2013

http://archive.org/details/universityofnort147univ

Measurement of Achievement in the Fundamental Elementary School Subjects

MEASURING THE SCHOOL PRODUCT

One of the most significant tendencies in education today is the stress which is coming to be laid on a definite knowledge of just what the school is producing. Most people who are engaged in educational work know in a general way that a given school system is well or poorly organized and administered, that its income is wastefully or economically expended, that it is or is not serving the needs of its community, that certain teachers are efficient, others not, that some pupils and classes are making satisfactory, others unsatisfactory, progress.

But school people, and the public at large, are rapidly becoming dissatisfied with merely general knowledge about such matters. Just as the business man, in the effort to make his business more efficient, employs experts, devises standards, analyses his records and those of his employees, the school is beginning to realize that judgments of school matters based on opinion must give way to judgments based on fact. Surveys of community and State systems, which attempt to judge the school product objectively, in terms of cold facts, have multiplied amazingly. The school is facing the question not only of whether on the whole it is doing good, but of just how efficiently it is making use of its resources in very definite ways.

The attempt to measure in an objective way the work of any system must involve the consideration of the system from many angles. And so we find that standards are being set by which to test the success of a system in holding its pupils, its expenditures, its service to its community, etc.

This bulletin is concerned with only one phase of the whole matter. It attempts to present briefly certain standard tests by which it is possible to measure results and progress in some aspects of a few fundamental subjects of the elementary school curriculum. The tests mentioned below are well known, have been widely used, and are almost all of proven value. Any competent teacher who carefully follows instructions can give and score these tests, and find in a definite way the standing of her pupils in the phase of work measured.

VALUE OF SUCH TESTS

Every teacher is obliged to pass judgment on the work of pupils by assigning definite marks in the various subjects. Why, it may be asked, are not teachers' marks satisfactory methods of estimating standing and

progress? Why is it necessary to substitute objective tests? The answer is simple. In the first place teachers' marks are far more unreliable than most teachers suppose. Not only has it been found that different teachers in schools of the same grade mark the same papers so differently that the marks given may range from far below the passing mark to approximate perfection, but the marks given by the same teachers to the same paper at different times vary enormously. A poor night's sleep, a headache, distracting thoughts—all such factors may greatly disturb results. Especially inexact for marking purposes is the common method of marking on a scale of one hundred per cent. There is probably no teacher in existence who can surely tell the difference between a paper that is worth 76 per cent and one that is worth 75 per cent. The single per cent is too small a unit—differences of 5 per cent are the smallest that can have any real meaning. Again, the standards used by teachers in different schools are so different that it is absolutely impossible to compare system with system, or to set up any general standards, based on teachers' marks, which pupils should reach. Marks are also very often poorly distributed among a class (see "normal distribution" below).

Such tests as those described below, on the other hand, are given and scored under perfectly definite conditions. The records which pupils make in these mean something more than the opinion of a single teacher. They can be compared with standards for the country as a whole.

One of the greatest values of the tests is that it is possible to measure by them the progress of a class during the year. The Courtis Tests in Arithmetic, for example, can be given in the fall and again toward the close of the year, and it can easily be noted whether the class as a whole, or individual pupils, have made satisfactory progress. Using these same tests, which measure the pupil's ability in the four fundamental operations of arithmetic, as an illustration, it will be seen that a comparison of the scores made in the four operations with standard scores devised from the records of many thousands of children all over the country in the same tests gives the teacher an exact basis on which to judge the standing of her class.

The value of such tests in enabling the superintendent to judge the efficiency of elementary teachers is obvious. The results which they give are not clouded by the many extraneous factors which often lead a superintendent astray in his judgments of his teachers.

But let it be emphasized that it is no function of the tests to "show up" a weak teacher and stop with that. The results from the tests are only the beginning, they should serve as a basis for constructive criticism of the teacher's work by herself and by the superintendent. The tests simply open up the field. If results are not satisfactory, what is wrong? It may not be the teachers. It may be that the methods of

teaching imposed by the texts in use, or by the traditions of the system, are at fault. *What* is wrong the progressive teacher or superintendent will set himself to determine. Teachers should welcome such tests, remembering that their purpose is to help them know the facts about their own work.

It is suggested, then, that the tests below be given by the teacher or superintendent near the beginning of the year's work, and again at the close, and the results compared. Compare your results, too, with the standard scores, and with those in other systems, but the determination of your own progress is most important. Examine not only the records of the class as a whole, but go over carefully the individual scores. Note particularly the pupils who show up poorly, and make up your mind to find out why, and what can be done to bring these pupils up to a higher level. The possibilities which the tests offer of individual diagnosis are not the least of their values.

Remember that there is such a thing as making too high a score in the tests. A class which does this may have gained efficiency in some one thing at the sacrifice of efficiency in others. Over-drill in the fundamentals of arithmetic, for example, is as bad as under-drill. True efficiency means a balanced product, not extraordinary feats in a few lines of work.

Caution. In giving the tests, follow directions to the letter. A change that seems immaterial to you may invalidate the whole result. For this reason it is safe to compare only in the most general way the results attained in one system with those attained in another. The directions, however trivial they may seem, have been formulated with extreme care by the authors of the tests, and must be followed if the results are to be worth while.

STATISTICAL TERMS

A few simple statistical terms are used in some of the tests, the meaning of which may require a little explanation. One of these is the *median.*

The median is a measure of the same general nature as the average. That is, it attempts to combine into one score the records of a group. It differs from the average in that it is the middle term of a series of figures when that series is arranged in order. To find the median, then, arrange the scores in either ascending or descending order and take the middle term. In case the series contains an even number of terms the median is the average of the two middle terms. Thus, consider the series 8, 9, 6, 4, 7, 3, 2. Arrange the series in order: 2, 3, 4, 6, 7, 8, 9. Since it contains seven terms, 6 is the median of the series. The median score of a class, then, is that score which half of the class reaches or exceeds. If your class contains 40 pupils, and the median score is 12, it

means that one-half the class made records as good as, or better than, 12, the other half not. (The statement above is roughly true; sufficiently so for all practical purposes.)

The median is more commonly used in educational measurements than the average for a number of reasons which need not be stated here.

Another measure in common use is the *average deviation.* It is the average amount by which individual scores differ from the median or average of the group. It is found by determining the amount by which the score of each individual differs from the median or average of the group, adding these differences (paying no attention to whether they are plus or minus), and dividing the sum by the number of individuals in the group. Thus, in the series 2, 3, 4, 6, 7, 8, 9, the individual variations from the median (6) are 4, 3, 2, 0, 1, 2, 3. Adding, and dividing by the number of terms in the series (7), we have 2.14 as the average deviation (A. D.).

It is evident that two series may have the same median or average, and yet have different average deviations. The more the individuals in the group vary from each other the greater will be the average deviation. A small average deviation means that the scores of the individuals are closely bunched around the median or average. Thus, two classes may both read with a median speed of 150 words per minute, but the pupils in one range from 50 to 300, in the other from 100 to 250. The former series evidently has the larger average deviation. A large average deviation, then, signifies great individual differences in achievement, and may mean, for example, that pupils are not properly graded. Just how great the average deviation should be in a given test is not a matter which can be covered by any rule; a comparison of the scores of various classes will give a basis of experience which will make it possible after a little experience to interpret the meaning of this measure.

A word is needed regarding the *normal distribution* of the scores of a group of children in a given test. If the group is large enough, and is homogeneous, the scores will be represented by a curve shaped like a bell; that is, the median or average score, and those near it on either side, will be the scores most frequently made, and the scores farther away from the median or average on either side will be fewer, and approximately the same in number on either side. Thus, suppose that a test is considered in which five scores, call them 1, 2, 3, 4, 5, are possible. If the group to which the tests are given is large and made up of children of approximately the same age, etc., the middle grade, or 3, will be attained by not far from 50 per cent of the group, grades 2 and 4 by about 22 per cent each; grades 1 and 5 by about 3 per cent each. The record made most frequently is called the *mode.* In a normal distribution the mode will also be the median and the average; the middle score will also be the most frequent.

It is of course not to be expected that a class of 40 pupils should show in any test exactly a normal distribution; the class ought, however, to approximate such a distribution in that the records made should be "bunched" around the average or median, and become fewer as the more extreme scores are reached. Consider the speed of reading, for example. If the median speed of the class is 150 words per minute, the individuals in the class should be bunched about this score. If one group is bunched about this, another about 75, another about 250, it means that there are in the class three natural groups, three modes, three levels of ability. Such a group is to all intents and purposes not a single grade in the subject tested, but three sections of varying ability. The natural argument is of course, if other tests reveal the same sort of grouping, the class is badly in need of division into sections.

In general, it must be remembered that it is not sufficient simply to give and score any standard test. Results must be studied and interpreted, and that interpretation utilized for constructive work.

For further reading. The literature dealing with tests and scales is widely scattered in technical journals. The only attempt to bring much of it together in a usable form is in the book by Daniel Starch, "Educational Measurements" (Macmillan, New York). This book should be read by every one who is interested in this phase of school efficiency. Lists of articles, etc., dealing with particular phases of the work will be furnished upon application to any member of the faculty of the School of Education.

ARITHMETIC

The Courtis Tests. "(Courtis Standard Research Tests, Arithmetic, Series B)." The Series makes a four-page folder, each page devoted to a series of examples in one of the four fundamental operations. A copy of the folder is needed for each pupil tested. These, with booklet of directions, forms for recording class scores, etc., cost thirty-five cents for a class of forty. A detailed "manual of instruction" for giving and scoring the tests costs seventy-five cents. All material may be obtained from Dr. S. A. Courtis, 82 Eliot St., Detroit, Michigan.

A specified time is allowed the pupils for each part of the test, and the score is recorded for each individual as the number of examples attempted during that period, and the number right. No partial credits are given. On the forms for recording class scores will be found standards for the various grades from grade 3 through grade 8. The standard median accuracy for each grade is 100 per cent (one-half of the pupils in each grade should complete correctly all examples attempted). These standard scores, both as to speed and accuracy, represent theoretically desirable conditions rather than those actually found in school practice.

Scores based on actual school tests will be found in the following table. These are derived by combining the achievements of many thou-

sands of pupils in all sorts of school systems, at the end of the school year. "Accuracy" in each case means the accuracy reached or exceeded by one-half the class. (Accuracy is determined by the relation of the number of examples attempted to the number correctly done.) When classes are divided into A. and B. sections, the scores for the two sections are combined in the table.

TABLE I

MEDIAN SCORES IN SPEED AND ACCURACY IN THE VARIOUS GRADES.

(From Bulletin No. 4, Department of Co-operative Research, 82 Eliot St., Detroit).

GRADE	ADDITION		SUBTRACTION		MULTIPLICATION		DIVISION	
	Speed	Acc.	Speed	Acc.	Speed	Acc.	Speed	Acc.
3	6.3	41	5.6	49	0.8	----------	0.6	----------
4	7.4	64	7.4	80	6.2	67	4.6	57
5	8.6	70	9.0	83	7.5	75	6.1	77
6	9.8	73	10.3	85	9.1	78	8.2	87
7	10.9	75	11.6	86	10.2	80	9.6	90
8	11.6	76	12.9	87	11.5	81	10.7	91

Thus, from the table above, a fifth grade, counting together the scores of A. and B. sections, should at the end of the year be able in the prescribed times to do the examples at such a rate that one-half of the class will do 8.6 or more in addition, the class as a whole working with an accuracy of 70 per cent, 9 in subtraction with an accuracy of 83 per cent, etc.

The following table shows the results obtained in one of the city systems of North Carolina in the spring of 1915. The scores for the various sections of each grade are combined.

TABLE II.

GRADE	ADDITION		SUBTRACTION		MULTIPLICATION		DIVISION	
	Speed	Acc.	Speed	Acc.	Speed	Acc.	Speed	Acc.
4	4.8	49	4.6	60	3.8	39	2.7	46
5	6.4	61	6.6	69	6.2	74	3.5	73
6	7.0	63	8.4	76	7.8	77	6.2	84
7	7.8	66	9.4	79	9.3	78	8.1	86
8	8.7	68	10.6	78	9.6	81	8.5	88

Comparing these with the scores in Table I it is evident that the pupils were below the standards for the country as a whole in both speed and accuracy. The data for "efficiency"—that is, the number of pupils in the various grades who measure up to the scores set as standards for those grades—are not given here. These are also somewhat lower than those found by Courtis.

It is clear that most of the pupils in the system represented by Table II need more drill on the fundamentals, and that this drill should be directed toward increasing both the speed and the accuracy of the work.

The arithmetic tests should be given at least twice during the year, in fall and spring, in order that the improvement of the class during the year may be ascertained. The records of the individual pupils should be carefully noted, and the work shaped in such a way as to emphasize those points which stand out as weak.

But remember that it is possible to over-drill on the fundamentals as well as to drill too little. A class which far exceeds the suggested standards may be in as unsatisfactory a condition as one far below them. Such a condition may mean that other phases of arithmetical work have been neglected.

Remember, also, that those pupils who are deficient in the tests need individual attention. Do not depend on class instruction to bring about an improvement in their work.

The Woody Scales. These are designed to measure ability in the four fundamental operations. They differ from the Courtis tests in that the problems in each series are of increasing difficulty, the first being very simple, thus permitting their use as far down as the second grade, while the last are difficult enough to be beyond the ability of half of the average eighth grade. Two series are available in each of the four fundamentals. Series B is an abridgment of Series A, the latter being of most value to teachers with some knowledge of the theories on which Educational Scales are based. Forms for these tests may be obtained from the Bureau of Publications, Teachers College, Columbia University, New York City, at a cost of 40 cents per 100 copies.

Problem-solving. Several measurements of the ability to solve arithmetical problems have been prepared. One of the most practicable scales is that devised by Starch. It consists of a series of problems of regularly increasing difficulty (the difficulty of the problems having been carefully determined by experiment). There are fifteen of these problems. The simplest is "Mary had 4 apples and her mother gave her 7 more. How many apples did Mary then have?" The most difficult is, "John had $1.20 Monday. He earned 30 cents each day on Tuesday, Wednesday, Thursday and Friday. Saturday morning he spent one-third of what he had earned in the four days. Saturday afternoon his father gave John half as much as John then had. How much did his father give John?"

For copies of this test, and information as to price, instructions for giving, etc., address Dr. Daniel Starch, University of Wisconsin, Madison, Wisconsin.

A second scale for the measurement of ability to solve arithmetical problems is C. W. Stone's "Reasoning Test in Arithmetic," published by the Bureau of Publications, Teachers College, Columbia. Price, 40 cents per 100 copies.

Warning. Do not assume that because pupils test well in the fundamentals they therefore possess necessarily a well-rounded knowledge of arithmetic. All the experimental work shows that arithmetical abilities are specialized, so that the possession of a given grade of ability in one sort of operation is by no means a sure indication of the pupil's ability in other phases of arithmetical work.

The Time Element in Arithmetic. Jessup reports in the "Fourteenth Yearbook of the National Society for the Study of Education" an investigation into the time devoted to arithmetic, according to the best current practices (as based on 630 cities). This should be about as follows:

The first grade not to exceed 75 minutes per week.

The second grade not to exceed 100 minutes per week.

The third grade not to exceed 125 minutes per week.

The fourth, fifth, sixth, and seventh grades not to exceed 150 minutes per week.

The eighth grade not to exceed 170 minutes per week.

The amount of this time given over to drill work is as follows:

First grade, 43 per cent of the total time given to arithmetic.

Second grade, 50 per cent of the total time given to arithmetic.

Third grade, 52 per cent of the total time given to arithmetic.

Fourth grade, 45 per cent of the total time given to arithmetic.

Fifth grade, 39 per cent of the total time given to arithmetic.

Sixth grade, 31 per cent of the total time given to arithmetic.

Seventh grade, 22 per cent of the total time given to arithmetic.

Eighth grade, 17 per cent of the total time given to arithmetic.

HANDWRITING

Several scales for the measurement of handwriting are available. Those most in use are "A Measuring Scale for Handwriting," by L. P. Ayres (Russell Sage Foundation, Division of Educational Research, New York City. Five cents per copy. A bulletin describing the method of deriving the scale is also five cents), and "A Scale for Handwriting of Children in Grades 5 to 8," by E. L. Thorndike. (Bureau of Publications, Teachers College, Columbia University, New York City, eight cents per copy.)

Each scale consists of a series of samples of handwriting of known excellence. The rank of samples in the Ayres Scale was determined by their legibility, that of those in the Thorndike Scale by their "general merit."

A more elaborate scale is that derived by F. N. Freeman (Houghton-Mifflin Co., Boston, Mass. Twenty-five cents per copy). In this scale the order of merit of the samples has been based on several characteristics. The scale should be of interest to teachers of handwriting.

The Ayres Scale is probably on the whole the most satisfactory for general use. In using the scale, secure samples of the handwriting of the pupils, written under natural conditions (that is, do *not* use samples written under conditions in which the pupils have been told to write "as well as possible" or "as speedily as possible." All pupils may be required to write the same sentence or stanza of poetry. "The sly fox jumped quickly over the lazy brown dog" is a sentence which contains all the letters of the alphabet. "Mary had a little lamb" has often been used in giving tests.

In order to determine the quality of a given sample, start at the lower end of the scale and compare the sample with the samples of the scale until a quality is reached on the scale which is distinctly better than the pupils' sample. Then start at the upper end of the scale, and compare the sample with those of the scale until a quality of distinctly inferior grade is reached on the scale. From a consideration of those two judgments it is usually possible to decide with little difficulty where the given sample belongs. Samples that seem, for example, better than 70 but not so good as 80 should be graded 70, as the grade 70 means from 70 up to 80. Freeman, writing in the Fourteenth Yearbook of the National Society for the Study of Education, sets the following standards for the various grades, in terms of the Ayres Scale:

The second grade should write with an average quality of 44 and an average speed of 36 letters per minute.

The third grade should write with an average quality of 47 and an average speed of 48 letters per minute.

The fourth grade should write with an average quality of 50 and an average speed of 56 letters per minute.

The fifth grade should write with an average quality of 55 and an average speed of 65 letters per minute.

The sixth grade should write with an average quality of 59 and an average speed of 72 letters per minute.

The seventh grade should write with an average quality of 64 and an average speed of 80 letters per minute.

The eighth grade should write with an average quality of 70 and an average speed of 90 letters per minute.

The above standards should be reached with an expenditure of time of not over 75 minutes per week. The class should be measured near the beginning of the year, and again near the close, and progress noted. It is suggested that pupils whose handwriting is of quality 80 or better on the Ayres Scale be excused from further practice. Remember that handwriting is meant to be read, and that continued drill after a high stand-

ard of legibility has been reached is a waste of the pupil's time for which there is, under modern conditions, no excuse.

The handwriting of pupils who seem unable to make satisfactory progress is best examined by the Freeman Scale, and in this connection his book, "The Teaching of Handwriting" (Houghton-Mifflin Co., Boston. Sixty cents), is valuable.

In view of the current interest in handwriting in the schools of the State, it is much to be desired that the results of the various methods in use should be compared. This may be done by taking two classes, or two sections of the same class, which are as nearly as possible alike in their ability at the start, and teaching one section by one method, one by another, for a given period of time, keeping all other conditions than the method used the same for both classes. At the end of the specified time the handwriting of the two classes may be graded by the Ayres or other scale, and comparisons made. Remember that not only the excellence of any method as such but the amount of time it needs to produce results must be considered in judging results. Any method which requires on the average more than fifteen minutes of the pupil's time per day should be carefully scrutinized, as the presumption is that it succeeds only at the cost of an undue amount of time and energy.

READING

Various processes involved in reading are capable of measurement. Simplest of all is naturally the measurement of the speed of reading. This may be determined for either silent or oral reading.

The Speed of Reading Silently. It is more important than is generally recognized that pupils should be taught to read silently at a fair rate of speed. The rapidity with which a class can read has an important bearing on the length of lessons which may be assigned in any subject, for example. The discovery that various individuals read with extreme slowness may sometimes show why work done by them is unsatisfactory. The discovery that the class scores as a whole are below the suggested standards means more drill on silent reading for the whole class.

Remember that it is silent reading, and not oral reading, of which the pupil will in after life have greatest need. Do you practice your pupils on this important phase of their work?

Each individual has at least two characteristic rates of silent reading; there is the rate at which we read novels, newspapers, magazines, fiction, etc.—call this the *normal* rate. There is, second, the rate at which we read material which is to be remembered; a lesson, for example. Call this the *careful* rate. It is evident that for school purposes it is the careful rate which it is most desirable to know. This may be tested by the following simple experiment:

Take a simple prose passage, such as is to be found in a third or fourth grade reader. Choose for this purpose a book of which there are copies enough in the school to put one into the hands of every member of the grade which is being tested. The reader in customary use is suitable, provided you are sure that the passage you select is a sight passage, and not familiar to the members of the class. Be sure that the passage is simple. A fourth grade reader will do to test all grades.

Be sure that the class does not begin to read the passage until told. Say, "I am going to ask you to read a story in (name of book). You will not have time to finish this story, for at the end of two minutes I am going to tell you to stop. You must stop with the word you are reading, even though it is in the middle of a sentence. Then with your pencil draw a line around the last word you read. Read carefully, for as soon as you have finished I am going to ask you to write all you can remember of what you read."

Have the class take books, paper and pen or pencil. Tell class the page on which the story begins, and ask them to turn to that page, but to be sure not to begin reading. To make this sure, let the book be closed as soon as the page is found, the pupil keeping his finger between the pages so that it can be opened immediately as you say begin. Be sure that every member of the class understands exactly what is to be done and is ready before you start the test. Give an opportunity for questions. When all is ready say, "Open books. Begin!" At the end of exactly two minutes say "Stop! Draw a line around the last word you read."

To determine the number of words read per minute it is simply necessary to count the total number of words read by each pupil and divide the result by two. The work may be shortened by counting the total number of words to the end of each line in your own copy of the book and recording the result for reference. Thus, if the first line of the passage contains eleven words, the second ten, the third twelve, etc., write after the first line in your book 11, after the second 21, after the third 33, etc. This will make it possible to determine the total number of words read by a pupil by simply locating the line which contains the last word he read, noting the number standing at the end of that line in your book and subtracting the words not read by the pupil in that line from the total.

The test will not usually be given below the fourth grade.

Courtis, in the Fourteenth Yearbook of the National Society for the Study of Education, sets the following standards for the various grades in careful reading. The scores are in terms of the median achievement of the class.

The fourth grade should read with a median speed of 110 words per minute.

The fifth grade should read with a median speed of 130 words per minute.

The sixth grade should read with a median speed of 160 words per minute.

The seventh grade should read with a median speed of 185 words per minute.

The eighth grade should read with a median speed of 200 words per minute.

The adult should read with a median speed of 200 words per minute.

The rate of oral reading can be measured by noting the time at which a given pupil begins reading aloud, the amount he reads in a given time, and the expression of this in words per minute. Oral reading is naturally slower than silent.

Note. Do not be afraid that pupils will read silently too rapidly. There is little danger of this. Results obtained in the Cleveland Survey, for example, show that those pupils who excel in silent reading also excel in the understanding of what they read, and vice versa. A slow rate of silent reading may be due to too much attention to the mechanical side of reading, as in the great over-emphasis on phonics which has for a number of years handicapped many schools in North Carolina, or to a wandering attention. In either case it is bad. Practice the class in silent reading.

Reproduction of Material Read. As a part of the test in silent reading it was required that the pupils should write as much as possible of what was read. The material so written may be utilized to test the pupils' power of reproducing material which has been read. It is evident that it makes little difference whether the pupil reproduces the exact words of the text read; it is his power to reproduce the ideas in which you are interested. In order to test his ability to do this it is necessary to divide the passage read into units, each of which expresses a single idea. For example, take the following passage, "One day Mr. Lincoln was out riding. As he passed along the road he saw a pig sinking into a mudhole." This may be divided into unit ideas as follows: One day : Mr. Lincoln : was out riding. : As he passed along the road : he saw a pig : sinking : into a mudhole.

Each word or phrase between the colons is a single idea. All that is necessary is that this idea—not the exact words in which it is expressed in the text—should be reproduced by the pupil. For example, if he says, "as he went along the road" or "passing along the road," the idea is correct. A little thought will make it possible to divide any passage that you may use into ideas in a similar fashion.

The exact number of ideas which should be reproduced naturally varies with many conditions. For practical purposes, provided the passage you have used is simple, pupils whose scores show a reproduction of forty to fifty or higher per cent of ideas may be regarded as satis-

factory. Those whose scores fall materially below this figure should be watched carefully in their class work to see whether the failure is due to carelessless or to inability, and measures taken accordingly.

Teachers wishing to measure more accurately the rate of reading and ability to reproduce material read than can be done by the use of passages from readers in common use in the school should write for copies of tests devised by Starch for this purpose. These consist of a series of paragraphs of uniformly increasing difficulty, together with instructions for scoring, etc. (Address Prof. Daniel Starch, University of Wisconsin, Madison, Wisconsin.)

Understanding of Material Read. The test of reproduction above is naturally also a test of the understanding of what was read. It is best, however, to test comprehension by a means that does not at the same time involve memory. The most widely used test for this purpose is that devised by Kelly, the "Kansas Silent Reading Test." (Write to the Bureau of Educational Measurements and Standards, State Normal School, Emporia, Kansas. A sheet will be needed for each pupil tested. Price of sheets and necessary record forms, thirty-three cents per hundred.) The test is simple to give and to score. It consists of a series of paragraphs of increasing difficulty, each paragraph having been assigned a value on the basis of experiment. Each paragraph tells the pupil to do something, write an answer, underline a word, etc., which he must understand the paragraph as a whole in order to do. Five minutes are allowed, the pupil to read as many paragraphs as possible and do what is indicated in each during the allotted time.

One test is to be used in the third, fourth and fifth grades, a second, somewhat more difficult, in the sixth, seventh and eighth, and a third in the ninth, tenth, eleventh and twelfth. Directions for the simple scoring necessary and tentative standard scores are given in the material which will be supplied you. The scoring simply consists in adding the values of the paragraphs answered correctly. No value is given unless the answer is correct.

It will be seen that such a test depends upon speed of reading as well as comprehension, and some of the questions asked are of such a nature as to be almost "catch-questions." The test, however, is a valuable one.

Thorndike has devised a scale for measuring the comprehension of words silently read, and one for measuring the comprehension of paragraphs. (Teachers' College Record, September, 1914.) No time limit is set for these tests. The test for the measuring of comprehension of words (Reading Scale A) consists of a series of lines of isolated words, each line containing words of equal difficulty, and the difficulty of the lines increasing by a uniform amount. These words the pupils are to examine and write a specified letter or word under each which has a given meaning, for example, the letter F under every word that means a flower, A under each meaning an animal, etc. The first line, for ex-

ample, is "camel, Samuel, bird, lily, cruel." The score is simply the value assigned to the highest step passed, "passing" being defined as understanding correctly four out of five of the given words.

A revision and extension of this scale is reported by Thorndike in the "Teachers' College Record" for November, 1916. The Scales are published in a convenient form for use. Price, 40 cents per 100.

Thorndike's Scale for measuring the understanding of sentences (Scale alpha) consists of a series of paragraphs of increasing difficulty, each followed by a series of questions to be answered after the paragraph has been read. The paragraphs may be re-read as often as the pupil desires. A revision of this Scale is published in a convenient form for use in testing. Price, 75 cents per 100 copies. Record sheets, 20 cents a dozen. Two forms will be needed, one for grades three to five, one for grades six to eight.

No standard scores are as yet available for these tests.

Tests by which both oral and silent reading may be measured were devised by W. S. Gray and used in the work of the Cleveland Survey. Both series consist of selections of increasing difficulty, and may be obtained from the author, School of Education, University of Chicago, Chicago, Ill. Each series, 50 cents per 100.

H. A. Brown has also published a silent reading test for use in the elementary school. (Bureau of Research, 25 Capitol Street, Concord, N. H. Also ask for bulletin on "The Measurement of Ability to Read.")

Language Scales. Not as tests of reading ability, but of what may roughly be termed language ability, the completion tests devised by Trabue are important. (Trabue, M. R., Completion-Test Language Scales. Write Bureau of Publications, Teachers' College, Columbia University, for printed forms for use in giving the tests.)

Each of the Scales consists of a series of sentences printed in incomplete form, as: "Time ..... often more valuable ..... money." The task is to complete the sentence, writing only one word in each blank. Each scale begins with sentences which are very simple to complete, and increases in difficulty by known amounts. A single scale may be used from the second or third grade through the high school. Completions are scored as correct (two points credit for each sentence perfectly completed), partially correct (one point credit for each sentence), or wrong (no credit). The method of scoring is thus very simple. The abilities which the Scales test are very closely related to the ability of pupils to deal with language, and to those abilities which determine the promotion or lack of promotion of an individual. "The completion test is not proposed as a substitute for the judgment of the teachers in promoting pupils, but rather as an aid to it. Employing one of these Scales will call attention to those children who have unusual ability to under-

stand and interpret printed words and phrases. These unusual cases may then be investigated more carefully and such readjustments made as will allow these children to work under the more favorable conditions."

SPELLING

The most commonly used Scale for measuring spelling ability is the Ayres "Measuring Scale for Ability in Spelling." (Russell Sage Foundation, Division of Education, 130 East 22d Street, New York City. Five cents.) The scale consists of one thousand words, which have a fair claim to being considered the most common words in English written language, as determined by a consideration of papers written by 70,000 children in 84 different cities. These words are arranged in columns, the words in each column being of equal difficulty, as determined by the percentage of children spelling them correctly, and the differences in difficulty between adjoining columns being approximately equal. At the top of each column are numbers indicating about what per cent of correct spellings may be expected of the children in a given grade. For example, over column K are the numbers 58, 79, 92, 96, 99. Standing opposite 58 is "second grade," opposite 79, "third grade," etc. This means that if a spelling test is given consisting of words in column K the average score of the entire second grade should be about 58 per cent, of the third grade, 79 per cent, the fourth grade, 92 per cent, etc.

The importance to the teacher of the knowledge to be derived from a use of the scale is obvious. Remember that the words a pupil will need to know how to spell are those which he will need to use in writing. There is no virtue in being able to spell words like "pneumogastric" and "phthysic," if one cannot spell correctly the common words of the Ayres Scale. There is no magical process which makes it possible for a pupil to spell common words which he has never studied because he is able to spell uncommon words which he has painfully studied. Words from the Ayres list should be used several times during the year for tests, to check up the spelling ability of the class.

Starch has devised a series of lists for testing spelling ability, each list consisting of one hundred words of different difficulty, the difficulty of any one of the lists as a whole being approximately equal to that of each of the others. The score is the number of words spelled correctly. These lists, with directions for giving and scoring the tests, and Standard Scores for the various grades through the eighth, are contained in the "Measurement of Efficiency in Reading, Writing, Spelling and English," by Daniel Starch, University of Wisconsin, Madison, Wis.

The Time Element. It is easily possible to give up too much of the total school day to spelling. The few data available as to the best cur-

rent practices in this respect, as summarized by Pryor in the Fourteenth Yearbook of the National Society for the Study of Education, indicate that about six per cent of the total time for all subjects is representative. The amount of time devoted to spelling has been gradually but surely reduced in the best systems.

Here, as elsewhere, it is always the first impulse of a teacher whose class does not come up to standard to increase the amount of time devoted to the subject. But careful studies have shown that there is, as a rule, little relationship between the amount of time devoted to a given subject and excellence in that subject. The logical conclusion is, *if results are unsatisfactory, try to improve methods of teaching* rather than increase the time given to the subject much beyond the best current standards.